

Adaptation of the Alicki-Fannes-Winter method for the set of states with bounded energy and its use.

M.E. Shirokov

Steklov Mathematical Institute, Moscow, Russia

Abstract

We describe a modification of the Alicki-Fannes-Winter method (used for proving uniform continuity of functions on the set of quantum states). It allows to show uniform continuity on the set of states with bounded energy of any approximately affine function having limited growth with increasing energy.

Some applications in quantum information theory are considered. In particular, the uniform finite-dimensional approximation theorem for the Holevo capacity of energy constrained channels is proved.

Contents

1	Introduction and preliminaries	2
2	Basic results	5
3	Applications	7
3.1	Continuity bound for linear combinations of marginal entropies under energy constraint	7
3.2	Continuity bound for the output Holevo quantity not depending on a channel	11
3.2.1	Discrete ensembles	11
3.2.2	Continuous ensembles	14
3.3	On uniform finite-dimensional approximation of the Holevo capacity of a channel with energy constraint.	15

1 Introduction and preliminaries

Alicki and Fannes obtained in [1] a continuity bound (estimate for variation) for the quantum conditional entropy by using the elegant geometric method. Recently Winter proposed modification of this method which makes it possible to derive a tight continuity bound for the conditional entropy [20]. In fact, this method (in what follows we will call it Alicki-Fannes-Winter method, briefly, AFW-method) is quite universal, it gives uniform continuity bound for any bounded function F on the set $\mathfrak{S}(\mathcal{H})$ of quantum states which is not "too convex and too concave" in the following sense

$$-a(p) \leq F(p\rho + (1-p)\sigma) - pF(\rho) - (1-p)F(\sigma) \leq b(p), \quad \rho, \sigma \in \mathfrak{S}(\mathcal{H}), p \in [0, 1],$$

where $a(p)$ and $b(p)$ are nonnegative functions on $[0, 1]$ vanishing as $p \rightarrow +0$. Functions F satisfying this condition will be called *approximately affine*.¹

In particular, the AFW-method shows that any approximately affine bounded function on $\mathfrak{S}(\mathcal{H})$ is uniformly continuous on $\mathfrak{S}(\mathcal{H})$.

The AFW-method can be used regardless of the dimension of the underlying Hilbert space \mathcal{H} under the condition that F is a bounded function on the whole set of states. But in analysis of infinite-dimensional quantum systems we often deal with functions which are bounded and approximately affine only on the sets of states with bounded energy, i.e. states ρ satisfying the inequality

$$\mathrm{Tr} H \rho \leq E, \tag{1}$$

where H is a positive operator – the Hamiltonian of a quantum system associated with the space \mathcal{H} [3, 4, 19, 20].

The main obstacle for direct application of the AFW-method to functions on the set of states with bounded energy consists in the difficulty to estimate the energy of the states proportional to the operators $[\rho - \sigma]_{\pm}$ for any states ρ and σ satisfying (1). In this paper we show that this problem can be solved by using simple modification of the AFW-method. The main idea of this modification is using the operators $\mathrm{Tr}_R [\hat{\rho} - \hat{\sigma}]_{\pm}$, where $\hat{\rho}$ and $\hat{\sigma}$ are appropriate purifications of given states ρ and σ satisfying (1).

The modified AFW-method makes it possible to obtain continuity bound for any approximately affine bounded function F on the set of states satisfying (1). This continuity bound implies uniform continuity of F provided

¹I would be grateful for any comments concerning this terminology.

that

$$\sup_{\text{Tr} H \rho \leq E} F(\rho) = o(\sqrt{E}), \quad E \rightarrow +\infty. \quad (2)$$

Condition (2) is essential (note that the affine function $\rho \mapsto \text{Tr} H \rho$ may be discontinuous on the set of states satisfying (1)). Fortunately, this condition is valid for many entropic characteristics of states of a quantum system provided the Hamiltonian H satisfies the condition

$$\lim_{\lambda \rightarrow +0} [\text{Tr} e^{-\lambda H}]^\lambda = 1,$$

which holds, in particular, for the system of quantum oscillators playing central role in continuous variable quantum information theory.

Let \mathcal{H} be a separable infinite-dimensional Hilbert space, $\mathfrak{B}(\mathcal{H})$ the algebra of all bounded operators with the operator norm $\|\cdot\|$ and $\mathfrak{T}(\mathcal{H})$ the Banach space of all trace-class operators in \mathcal{H} with the trace norm $\|\cdot\|_1$. Let $\mathfrak{S}(\mathcal{H})$ be the set of quantum states (positive operators in $\mathfrak{T}(\mathcal{H})$ with unit trace) [3, 10, 18].

Denote by $I_{\mathcal{H}}$ the identity operator in a Hilbert space \mathcal{H} and by $\text{Id}_{\mathcal{H}}$ the identity transformation of the Banach space $\mathfrak{T}(\mathcal{H})$.

If quantum systems A and B are described by Hilbert spaces \mathcal{H}_A and \mathcal{H}_B then the bipartite system AB is described by the tensor product of these spaces, i.e. $\mathcal{H}_{AB} \doteq \mathcal{H}_A \otimes \mathcal{H}_B$. A state in $\mathfrak{S}(\mathcal{H}_{AB})$ is denoted ω_{AB} , its marginal states $\text{Tr}_{\mathcal{H}_B} \omega_{AB}$ and $\text{Tr}_{\mathcal{H}_A} \omega_{AB}$ are denoted respectively ω_A and ω_B .

The *von Neumann entropy* $H(\rho) = \text{Tr} \eta(\rho)$ of a state $\rho \in \mathfrak{S}(\mathcal{H})$, where $\eta(x) = -x \log x$, is a concave nonnegative lower semicontinuous function on the set $\mathfrak{S}(\mathcal{H})$ [3, 7, 18]. The concavity of the von Neumann entropy is supplemented by the inequality

$$H(p\rho + (1-p)\sigma) \leq pH(\rho) + (1-p)H(\sigma) + h_2(p), \quad (3)$$

where $h_2(p) = \eta(p) + \eta(1-p)$, valid for any states $\rho, \sigma \in \mathfrak{S}(\mathcal{H})$ and $p \in (0, 1)$.

The *quantum conditional entropy*

$$H(A|B)_\omega = H(\omega_{AB}) - H(\omega_B) \quad (4)$$

of a bipartite state ω_{AB} with finite marginal entropies is essentially used in analysis of quantum systems [3, 18]. The function $\omega_{AB} \mapsto H(A|B)_\omega$ is

continuous on $\mathfrak{S}(\mathcal{H}_{AB})$ if and only if $\dim \mathcal{H}_A < +\infty$.²

The conditional entropy is concave and satisfies the following inequality

$$H(A|B)_{p\rho+(1-p)\sigma} \leq pH(A|B)_\rho + (1-p)H(A|B)_\sigma + h_2(p) \quad (5)$$

for any $p \in (0, 1)$ and any states ρ_{AB} and σ_{AB} . Inequality (5) follows from concavity of the entropy and inequality (3).

The *quantum relative entropy* for two states ρ and σ in $\mathfrak{S}(\mathcal{H})$ is defined by the formula

$$H(\rho \parallel \sigma) = \sum \langle i | \rho \log \rho - \rho \log \sigma | i \rangle,$$

where $\{|i\rangle\}$ is the orthonormal basis of eigenvectors of the state ρ and it is assumed that $H(\rho \parallel \sigma) = +\infty$ if $\text{supp} \rho$ is not contained in $\text{supp} \sigma$ [3, 7].

The *quantum mutual information* of a state ω_{AB} of a bipartite quantum system is defined as follows

$$I(A:B)_\omega = H(\omega_{AB} \parallel \omega_A \otimes \omega_B) = H(\omega_A) + H(\omega_B) - H(\omega_{AB}), \quad (6)$$

where the second expression is valid if $H(\omega_{AB})$ is finite [8, 18].

Basic properties of the relative entropy show that $\omega \mapsto I(A:B)_\omega$ is a lower semicontinuous function on the set $\mathfrak{S}(\mathcal{H}_{AB})$ taking values in $[0, +\infty]$. It is well known that

$$I(A:B)_\omega \leq 2 \min \{H(\omega_A), H(\omega_B)\} \quad (7)$$

for any state ω_{AB} [9, 18].

The quantum mutual information is not concave or convex but the inequality

$$|pI(A:B)_\rho + (1-p)I(A:B)_\sigma - I(A:B)_{p\rho+(1-p)\sigma}| \leq h_2(p) \quad (8)$$

holds for $p \in (0, 1)$ and any states ρ_{AB} , σ_{AB} with finite $I(A:B)_\rho$, $I(A:B)_\sigma$. If ρ_{AB} , σ_{AB} are states with finite marginal entropies then (8) can be easily proved by noting that

$$I(A:B)_\omega = H(\omega_A) - H(A|B)_\omega, \quad (9)$$

and by using the concavity of the entropy and of the conditional entropy along with the inequalities (3) and (5). The validity of inequality (8) for any states ρ_{AB} , σ_{AB} with finite mutual information can be proved by approximation (using Theorem 1 in [14]).

²If $\dim \mathcal{H}_A < +\infty$ and $\dim \mathcal{H}_B = +\infty$ then formula (4) is not well defined for some states, but there is an alternative expression for $H(A|B)_\omega$ (derived from the below formula (9)) which gives concave continuous function on $\mathfrak{S}(\mathcal{H}_{AB})$ in this case [6].

2 Basic results

Let H be a positive operator in a Hilbert space \mathcal{H} and $E \geq E_0 \doteq \sup_{\|\varphi\|=1} \langle \varphi | H | \varphi \rangle$.

Then

$$\mathfrak{C}_{H,E} = \{\rho \in \mathfrak{S}(\mathcal{H}) \mid \text{Tr} H \rho \leq E\}$$

is a closed convex subset of $\mathfrak{S}(\mathcal{H})$. If H is the Hamiltonian of a quantum system associated with the space \mathcal{H} then $\mathfrak{C}_{H,E}$ is the set of states with mean energy not exceeding E .

Let F be a function defined on the set $\mathfrak{C}_{H,\infty} \doteq \bigcup_{E \geq E_0} \mathfrak{C}_{H,E}$. We will say that the function F is *approximately affine* if

$$-a(p) \leq F(p\rho + (1-p)\sigma) - pF(\rho) - (1-p)F(\sigma) \leq b(p) \quad (10)$$

for any $p \in [0, 1]$ and all $\rho, \sigma \in \mathfrak{C}_{H,\infty}$, where $a(p)$ and $b(p)$ are nonnegative functions on $[0, 1]$ vanishing as $p \rightarrow +0$.

Theorem 1. *If F is a function on $\mathfrak{C}_{H,\infty}$ possessing property (10) such that $B_F(E) \doteq \sup_{\rho \in \mathfrak{C}_{H,E}} |F(\rho)| < +\infty$ for all $E \geq E_0$ then*

$$|F(\rho) - F(\sigma)| \leq 2\sqrt{2\varepsilon} B_F(E/\varepsilon) + (1 + \sqrt{2\varepsilon})(a(\eta) + b(\eta)) \quad (11)$$

for any states ρ and σ in $\mathfrak{C}_{H,E}$ such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq \frac{1}{2}$, where $\eta = \sqrt{2\varepsilon}/(1 + \sqrt{2\varepsilon})$. The term $2B_F(E/\varepsilon)$ in (11) can be replaced by $B_F^+(E/\varepsilon) + B_F^-(E/\varepsilon)$, where $B_F^\pm(E/\varepsilon) \doteq \sup_{\rho \in \mathfrak{C}_{H,E}} \max\{\pm F(\rho), 0\}$.

Corollary 1. *If F is an approximately affine function on $\mathfrak{C}_{H,\infty}$ such that $B_F(E) = o(\sqrt{E})$ for $E \rightarrow +\infty$ then F is uniformly continuous on the set $\mathfrak{C}_{H,E}$ for any $E \geq E_0$.*

Proof of Theorem 1. Let $\mathcal{H}_R \cong \mathcal{H}$ and $\hat{\rho} = |\varphi\rangle\langle\varphi|$, $\hat{\sigma} = |\psi\rangle\langle\psi|$ be purifications of the states ρ and σ in $\mathcal{H} \otimes \mathcal{H}_R$ such that $\delta \doteq \frac{1}{2}\|\hat{\rho} - \hat{\sigma}\|_1 = \sqrt{2\varepsilon}$. Note that $\delta = \sqrt{1 - |\langle\varphi|\psi\rangle|^2}$.

Following [20] introduce the quantum states $\hat{\tau}_+ = \delta^{-1}[\hat{\rho} - \hat{\sigma}]_+$ and $\hat{\tau}_- = \delta^{-1}[\hat{\rho} - \hat{\sigma}]_-$ such that

$$\frac{1}{1+\delta} \hat{\rho} + \frac{\delta}{1+\delta} \hat{\tau}_- = \omega_* = \frac{1}{1+\delta} \hat{\sigma} + \frac{\delta}{1+\delta} \hat{\tau}_+.$$

By taking partial trace we obtain

$$\frac{1}{1+\delta} \rho + \frac{\delta}{1+\delta} \tau_- = \text{Tr}_R \omega_* = \frac{1}{1+\delta} \sigma + \frac{\delta}{1+\delta} \tau_+, \quad (12)$$

where $\tau_{\pm} = \text{Tr}_R \hat{\tau}_{\pm}$.

By using spectral decomposition of the operator $\hat{\rho} - \hat{\sigma} = |\varphi\rangle\langle\varphi| - |\psi\rangle\langle\psi|$ one can show that $\hat{\tau}_{\pm}$ are pure states corresponding to the unit vectors

$$|\gamma_{\pm}\rangle = p_{\pm}|\varphi\rangle + q_{\pm}|\psi\rangle, \text{ where } p_{\pm} = \frac{\langle\varphi|\psi\rangle}{\delta\sqrt{2(1\mp\delta)}}, \quad q_{\pm} = -\frac{(1\mp\delta)}{\delta\sqrt{2(1\mp\delta)}}.$$

So, we have

$$\begin{aligned} \text{Tr} H \tau_{\pm} &= \langle\gamma_{\pm}|H \otimes I_R|\gamma_{\pm}\rangle = |p_{\pm}|^2\langle\varphi|H \otimes I_R|\varphi\rangle + |q_{\pm}|^2\langle\psi|H \otimes I_R|\psi\rangle \\ &+ 2\Re \bar{p}_{\pm}q_{\pm}\langle\varphi|H \otimes I_R|\psi\rangle \leq |p_{\pm}|^2\text{Tr} H \rho + |q_{\pm}|^2\text{Tr} H \sigma + 2|p_{\pm}q_{\pm}|\sqrt{\text{Tr} H \rho}\sqrt{\text{Tr} H \sigma} \\ &\leq E(|p_{\pm}| + |q_{\pm}|)^2 = (1 + |\langle\varphi|\psi\rangle|)E/\delta^2 \leq 2E/\delta^2 = E/\varepsilon, \end{aligned}$$

where the Schwarz inequality was used.

It follows that the states τ_{\pm} belong to the set $\mathfrak{C}_{H,E/\varepsilon}$ and hence

$$|F(\tau_{\pm})| \leq B_F(E/\varepsilon). \quad (13)$$

By applying (10) to the convex decompositions in (12) we obtain

$$(1-p)[F(\rho) - F(\sigma)] \leq p[F(\tau_+) - F(\tau_-)] + a(p) + b(p)$$

and

$$(1-p)[F(\sigma) - F(\rho)] \leq p[F(\tau_-) - F(\tau_+)] + a(p) + b(p)$$

where $p = \frac{\delta}{1+\delta}$. These inequalities and upper bound (13) imply (11). The last assertion of the proposition follows directly from the above arguments, since $|F(\tau_+) - F(\tau_-)|$ is upper bounded by $B_F^+(E/\varepsilon) + B_F^-(E/\varepsilon)$. \square

Remark 1. In applications we often deal with a function F which is defined and approximately affine on the set $\mathfrak{C}_{H,\infty}^0 \doteq \bigcup_E \mathfrak{C}_{H,E}^0$, where $\mathfrak{C}_{H,E}^0$ is a convex subset of $\mathfrak{C}_{H,E}$ for each E (for example, $\mathfrak{C}_{H,E}^0$ is the subset of $\mathfrak{C}_{H,E}$ consisting of finite rank states, etc.). The proof of Theorem 1 shows that its assertion is valid for $\mathfrak{C}_{H,E}^0$ instead of $\mathfrak{C}_{H,E}$ if the following condition holds:

$$\text{the states } c_{\pm} \text{Tr}_R [\hat{\rho} - \hat{\sigma}]_{\pm} \text{ belong to the set } \mathfrak{C}_{H,\infty}^0, \quad (14)$$

where $c_{\pm} = \text{Tr}[\hat{\rho} - \hat{\sigma}]_{\pm}$, for any purifications $\hat{\rho}$ and $\hat{\sigma}$ in $\mathfrak{S}(\mathcal{H} \otimes \mathcal{H}_R)$ of arbitrary states ρ and σ in $\mathfrak{C}_{H,E}^0$.

Corollary 2. *Let $\mathfrak{C}_{H,E}^0$ be a dense subset of $\mathfrak{C}_{H,E}$ for each $E \geq E_0$ such that condition (14) holds. If F is an approximately affine function on $\mathfrak{C}_{H,\infty}^0$ such that*

$$B_F(E) \doteq \sup_{\rho \in \mathfrak{C}_{H,E}^0} |F(\rho)| = o(\sqrt{E}) \text{ for } E \rightarrow +\infty$$

then F has a uniformly continuous approximately affine extension \hat{F} to the set $\mathfrak{C}_{H,E}$ for any $E \geq E_0$ satisfying (11).

3 Applications

3.1 Continuity bound for linear combinations of marginal entropies under energy constraint

Several important entropic characteristics of a state of a finite-dimensional n -partite system $A_1 \dots A_n$ are defined as a real linear combination of marginal entropies, i.e. as a function

$$F(\omega_{A_1 \dots A_n}) = \sum_k \alpha_k H(\omega_{X_k}) \quad (15)$$

on the set of all states of the system, where ω_{X_k} is the partial state of $\omega_{A_1 \dots A_n}$ corresponding to the subsystem X_k of $A_1 \dots A_n$ and $\alpha_k \in \mathbb{R}$.

By using concavity of the von Neumann entropy and inequality (3) it is easy to show that the function F in (15) satisfies the following approximately affinity property

$$-a_F h_2(p) \leq F(p\rho + (1-p)\sigma) - pF(\rho) - (1-p)F(\sigma) \leq b_F h_2(p) \quad (16)$$

for all $p \in [0, 1]$ and any states $\rho, \sigma \in \mathfrak{S}(\mathcal{H}_{A_1 \dots A_n})$, where $a_F \leq \sum_{k: \alpha_k < 0} |\alpha_k|$ and $b_F \leq \sum_{k: \alpha_k > 0} \alpha_k$.³

It is also essential that many important characteristics F having form (15) possess lower and upper estimates proportional to one of the marginal entropies, i.e. they satisfy the inequality

$$-c_F^- H(\omega_B) \leq F(\omega_{A_1 \dots A_n}) \leq c_F^+ H(\omega_B), \quad (17)$$

³Inequality (8) shows that the coefficients a_F and b_F may be less than $\sum_{k: \alpha_k < 0} |\alpha_k|$ and $\sum_{k: \alpha_k > 0} \alpha_k$.

where B is a particular subsystem of $A_1 \dots A_n$ and c_F^-, c_F^+ are nonnegative numbers. For example, the quantum mutual information $I(A_1 : A_2)_\omega$ considered as a function of a state $\omega_{A_1 A_2 A_3}$ is nonnegative and upper bounded by one of the quantities:

$$2H(\omega_{A_1}), 2H(\omega_{A_2}), 2H(\omega_{A_1 A_3}), 2H(\omega_{A_2 A_3}).$$

In finite dimensions the properties (16) and (17) make it possible to directly apply the AFW-method to the function F and obtain the continuity bound

$$|F(\rho) - F(\sigma)| \leq (c_F^- + c_F^+) \varepsilon \log \dim \mathcal{H}_B + (a_F + b_F) g(\varepsilon), \quad (18)$$

where $\varepsilon = \frac{1}{2} \|\rho - \sigma\|_1$ and $g(\varepsilon) \doteq (1 + \varepsilon) h_2\left(\frac{\varepsilon}{1+\varepsilon}\right)$ [14, Proposition 1].

In infinite dimensions the function F in (15) is correctly defined if all the marginal entropies $H(\omega_{X_k})$ are finite (or at least the linear combination in (15) does not contain the uncertainty " $\infty - \infty$ "). So, the following problems naturally appear (cf.[14]):

- how to extend such narrow domain of definition of F ?
- how to analyse continuity properties of F ?

Solutions of the last problem for the entropy $H(\omega_{A_1})$ and for the conditional entropy $H(A_1|A_2)_\omega \doteq H(\omega_{A_1 A_2}) - H(\omega_{A_2})$ were recently proposed by Winter in [20], who obtained asymptotically tight continuity bounds for these quantities under the energy constraint on ω_{A_1} , i.e. the constraint defined by the inequality

$$\text{Tr} H_{A_1} \omega_{A_1} \leq E,$$

where H_{A_1} is the Hamiltonian of system A_1 satisfying the condition

$$\text{Tr} e^{-\lambda H_{A_1}} < +\infty \quad \text{for all } \lambda > 0. \quad (19)$$

In the case of $H(A_1|A_2)_\omega$ the role of system B in (17) is played by A_1 , since it is well known that $|H(A_1|A_2)_\omega| \leq H(\omega_{A_1})$ [3, 6, 18].

Winter's approach is based on combination of Fannes' type continuity bound (i.e. continuity bound of the form (18)) with special finite-dimensional approximation of arbitrary states with bounded energy. Application of this approach to any function F in (15) satisfying (16) and (17) is limited by the approximation step, since it requires special estimates depending on F . In contrast to this, the modified AFW-method (described in Section 2) makes

it possible to obtain an universal continuity bound for F under the energy constraint on the partial state ω_B corresponding to the system B in (17).

To formulate the main result of this section note that condition (19) with H_B instead of H_{A_1} implies that

$$\sup_{\text{Tr} H_B \rho \leq E} H(\rho) = H(\gamma_B(E)) \quad \text{for any } E > E_0 \doteq \sup_{\|\varphi\|=1} \langle \varphi | H_B | \varphi \rangle,$$

where $\gamma_B(E) \doteq e^{-\lambda(E)H_B} / \text{Tr} e^{-\lambda(E)H_B}$ is the *Gibbs state* of the system B corresponding to the energy E (the parameter $\lambda(E)$ is determined by the equality $\text{Tr} H_B e^{-\lambda H_B} = E \text{Tr} e^{-\lambda H_B}$) [17].

Proposition 1. *Assume that the inequalities (16) and (17) hold for the function F in (15) for all states with finite rank marginals.⁴ If the Hamiltonian H_B of the system B satisfies the condition*

$$\lim_{\lambda \rightarrow +0} [\text{Tr} e^{-\lambda H_B}]^\lambda = 1 \quad (20)$$

then $H(\gamma_B(E)) = o(\sqrt{E})$, $E \rightarrow +\infty$, and for any $E > E_0$ there exists a unique uniformly continuous extension \widehat{F} of the function F to the set

$$\mathfrak{C}_{H_B, E}^B = \{\omega_{A_1 \dots A_n} \mid \text{Tr} H_B \omega_B \leq E\} \quad (21)$$

such that

$$|\widehat{F}(\rho) - \widehat{F}(\sigma)| \leq (c_F^- + c_F^+) \sqrt{2\varepsilon} H(\gamma_B(E/\varepsilon)) + (a_F + b_F) g(\sqrt{2\varepsilon}), \quad (22)$$

for any states ρ and σ in $\mathfrak{C}_{H_B, E}^B$ such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq \frac{1}{2}$, where $g(x) \doteq (1+x)h_2(\frac{x}{1+x})$.

Condition (20) holds if the Hamiltonian H_B has the discrete spectrum $\{E_k\}_{k \geq 0}$ such that $\liminf_{k \rightarrow \infty} E_k / \log^q k > 0$ for some $q > 2$.⁵

Remark 2. Since condition (20) implies $H(\gamma_B(E)) = o(\sqrt{E})$, $E \rightarrow +\infty$, it guarantees that the main term in (22) tends to zero as $\varepsilon \rightarrow 0$.

Condition (20) is stronger than condition (19) with $A_1 = B$ which implies $H(\gamma_B(E)) = o(E)$, $E \rightarrow +\infty$ [13, Pr.1]. In terms of the sequence $\{E_k\}$ of eigenvalues of H_B condition (19) means that $\lim_k E_k / \log k = +\infty$. Hence,

⁴i.e. such states $\omega_{A_1 \dots A_n}$ that $\text{rank} \omega_{A_k} < +\infty$ for all $k = \overline{1, n}$.

⁵Lemma 3 in the Appendix shows that condition (20) is not valid if $\limsup_{k \rightarrow \infty} E_k / \log^2 k < +\infty$.

the last assertion of Proposition 1 shows that the difference between conditions (19) and (20) is not too large. It is essential that condition (20) holds for the Hamiltonian of the system of quantum oscillators [3, 19, 20].

Proof. Let $\bar{B} = A_1 \dots A_n \setminus B$ and $\hat{H} = H_B \otimes I_{\bar{B}}$ be a positive operator in $\mathcal{H}_{A_1 \dots A_n}$. In terms of Section 2 the set $\mathfrak{C}_{H_B, E}^B$ in (21) is $\mathfrak{C}_{\hat{H}, E}$. Let $\mathfrak{C}_{\hat{H}, E}^0$ be the subset of $\mathfrak{C}_{\hat{H}, E}$ consisting of states $\omega_{A_1 \dots A_n}$ such that $\text{rank} \omega_{A_k} < +\infty$ for all $k = \overline{1, n}$. It is easy to see that the family of subsets $\mathfrak{C}_{\hat{H}, E}^0$ satisfies condition (14). So, the main assertion of the proposition follows from Corollary 2 and Lemma 2 in the Appendix.

The last assertion of the proposition follows from Lemma 3 in the Appendix, since it is easy to see that

$$\left[\sum_{k=0}^{+\infty} e^{-\lambda E_k} \right]^\lambda = 1 \quad \Leftrightarrow \quad \left[\sum_{k=n}^{+\infty} e^{-\lambda E_k} \right]^\lambda = 1$$

for any sequence $\{E_k\}$ of positive numbers and any given n . \square

By applying Proposition 1 to the entropy and to the conditional entropy we obtain the following continuity bounds

$$|H(\rho_A) - H(\sigma_A)| \leq \sqrt{2\varepsilon} H(\gamma_A(E/\varepsilon)) + g(\sqrt{2\varepsilon}) \quad (23)$$

and

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 2\sqrt{2\varepsilon} H(\gamma_A(E/\varepsilon)) + g(\sqrt{2\varepsilon}) \quad (24)$$

under the condition $\text{Tr} H_A \rho_A, \text{Tr} H_A \sigma_A \leq E$, where $\varepsilon = \frac{1}{2} \|\rho - \sigma\|_1 \leq \frac{1}{2}$. These continuity bounds give more coarse estimates for variations than Winter's continuity bounds for these quantities obtained in [20]. This is not surprising, since Winter's method does not use the purifications of initial states implying appearance of the factor $\sqrt{\varepsilon}$ in (23) and (24).

The main advantage of continuity bound (22) is its universality. It allows to obtain continuity bounds under different forms of energy constrains. For example, by considering the mutual information $I(A:B)$ as a function on the set $\mathfrak{S}(\mathcal{H}_{ABC})$ and by using the inequality $0 \leq I(A:B) \leq I(A:BC)$, upper bound (7) and inequality (8) we obtain from Proposition 1 the following

Corollary 3. *If the Hamiltonian H_{BC} of subsystem BC of a tripartite system ABC satisfies condition (20) then the function $\omega_{ABC} \mapsto I(A:B)_\omega$ is uniformly continuous on the set of states with bounded energy of ω_{BC} and*

$$|I(A:B)_\rho - I(A:B)_\sigma| \leq 2\sqrt{2\varepsilon} H(\gamma_{BC}(E/\varepsilon)) + 2g(\sqrt{2\varepsilon}) \quad (25)$$

for any states ρ_{ABC} and σ_{ABC} such that $\text{Tr} H_{BC} \rho_{BC}, \text{Tr} H_{BC} \sigma_{BC} \leq E$ and $\frac{1}{2} \|\rho - \sigma\|_1 \leq \varepsilon \leq \frac{1}{2}$, where γ_{BC} is the Gibbs state of the system BC .

By using the Stinespring representation of a quantum channel it is easy to derive from Corollary 3 the following

Corollary 4. *Let $\Phi : A \rightarrow B$ be an arbitrary quantum channel and C be any system. If the Hamiltonian H_A of input system A satisfies condition (20) then the function $\rho_{AC} \mapsto I(B:C)_{\Phi \otimes \text{Id}_C(\rho)}$ is uniformly continuous on the set of states with bounded energy of ρ_A and*

$$|I(B:C)_{\Phi \otimes \text{Id}_C(\rho)} - I(B:C)_{\Phi \otimes \text{Id}_C(\sigma)}| \leq 2\sqrt{2\varepsilon} H(\gamma_A(E/\varepsilon)) + 2g(\sqrt{2\varepsilon}) \quad (26)$$

for any states ρ and σ in $\mathfrak{S}(\mathcal{H}_{AC})$ such that $\text{Tr} H_A \rho_A, \text{Tr} H_A \sigma_A \leq E$ and $\frac{1}{2} \|\rho - \sigma\|_1 \leq \varepsilon \leq \frac{1}{2}$, where γ_A is the Gibbs state of the system A .

The main term in (26) tends to zero as $\varepsilon \rightarrow 0$, since condition (20) implies $H(\gamma_A(E)) = o(\sqrt{E})$, $E \rightarrow +\infty$ (by Lemma 2 in the Appendix).

It is essential for applications that continuity bound (26) does not depend on the channel Φ . This will be used in the next section.

3.2 Continuity bound for the output Holevo quantity not depending on a channel

3.2.1 Discrete ensembles

Corollary 4 can be used for analysis of continuity properties of the output Holevo quantity

$$\chi(\{p_i, \Phi(\rho_i)\}) \doteq \sum_i p_i H(\Phi(\rho_i) \| \Phi(\bar{\rho})), \quad \bar{\rho} = \sum_i p_i \rho_i,$$

of a given channel $\Phi : A \rightarrow B$ with respect to variations of input *discrete ensemble* $\{p_i, \rho_i\}$ – a finite or countable collection $\{\rho_i\}$ of input states with the corresponding probability distribution $\{p_i\}$.

We will use three different measures of divergence between discrete ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$. The quantity

$$D_0(\mu, \nu) \doteq \frac{1}{2} \sum_i \|p_i \rho_i - q_i \sigma_i\|_1 \quad (27)$$

is a true metric on the set of all ensembles of quantum states considered as *ordered* collections of states with the corresponding probability distributions. It coincides (up to the factor 1/2) with the trace norm of the difference between the corresponding *qc*-states $\sum_i p_i \rho_i \otimes |i\rangle\langle i|$ and $\sum_i q_i \sigma_i \otimes |i\rangle\langle i|$ [18].

The main advantage of D_0 is a direct computability, but from the quantum information point of view we have to consider an ensemble of quantum states $\{p_i, \rho_i\}$ as a discrete probability measure $\sum_i p_i \delta(\rho_i)$ on the set $\mathfrak{S}(\mathcal{H})$ (where $\delta(\rho)$ is the Dirac measure concentrating at a state ρ) rather than ordered (or disordered) collection of states. If we want to identify ensembles corresponding to the same probability measure then it is natural to use the factorization of D_0 , i.e. the quantity

$$D_*(\mu, \nu) \doteq \inf_{\mu' \in \mathcal{E}(\mu), \nu' \in \mathcal{E}(\nu)} D_0(\mu', \nu') \quad (28)$$

as a measure of divergence between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$, where $\mathcal{E}(\mu)$ and $\mathcal{E}(\nu)$ are the sets of all countable ensembles corresponding to the measures $\sum_i p_i \delta(\rho_i)$ and $\sum_i q_i \delta(\sigma_i)$ respectively.

It is shown in [15] that the factor-metric D_* coincides with the EHS-distance D_{ehs} between ensembles of quantum states proposed by Oreshkov and Calsamiglia in [11] and that D_* generates the weak convergence topology on the set of all ensembles (considered as probability measures).⁶

The metric $D_* = D_{\text{ehs}}$ is more adequate for continuity analysis of the Holevo quantity, but difficult to compute in general.⁷ It is clear that

$$D_*(\mu, \nu) \leq D_0(\mu, \nu) \quad (29)$$

for any ensembles μ and ν . But in some cases the metrics D_0 and D_* are close to each other or even coincide. This holds, for example, if we consider small perturbations of states or probabilities of a given ensemble.

The third useful metric is the Kantorovich distance

$$D_K(\mu, \nu) = \frac{1}{2} \inf_{\{P_{ij}\}} \sum P_{ij} \|\rho_i - \sigma_j\|_1 \quad (30)$$

between ensembles $\mu = \{p_i, \rho_i\}$ and $\nu = \{q_i, \sigma_i\}$, where the infimum is over all joint probability distributions $\{P_{ij}\}$ with the marginals $\{p_i\}$ and $\{q_i\}$, i.e.

⁶This means that a sequence $\{\{p_i^n, \rho_i^n\}\}_n$ converges to an ensemble $\{p_i^0, \rho_i^0\}$ with respect to the metric D_* if and only if $\lim_{n \rightarrow \infty} \sum_i p_i^n f(\rho_i^n) = \sum_i p_i^0 f(\rho_i^0)$ for any continuous bounded function f on $\mathfrak{S}(\mathcal{H})$.

⁷For finite ensembles it can be calculated by a linear programming procedure [11].

such that $\sum_j P_{ij} = p_i$ for all i and $\sum_i P_{ij} = q_j$ for all j . Since $D_* = D_{\text{ehs}}$, it is easy to show (see [11]) that

$$D_*(\mu, \nu) \leq D_K(\mu, \nu) \quad (31)$$

for any discrete ensembles μ and ν .

For our aims it is essential that the Kantorovich distance has natural extension to the set of all generalized (continuous) ensembles which generates the weak convergence topology on this set (see the next subsection).

In the following proposition we assume that the set of all discrete ensembles is equipped with the weak convergence topology (generated by the metrics D_* and D_K).

Proposition 2. *Let $\Phi : A \rightarrow B$ be an arbitrary quantum channel. If the Hamiltonian H_A of input system A satisfies condition (20) then the function $\{p_i, \rho_i\} \rightarrow \chi(\{p_i, \Phi(\rho_i)\})$ is uniformly continuous on the set of all ensembles $\{p_i, \rho_i\}$ with bounded average energy $E(\{p_i, \rho_i\}) \doteq \sum_i p_i \text{Tr} H_A \rho_i$ and*

$$|\chi(\{p_i, \Phi(\rho_i)\}) - \chi(\{q_i, \Phi(\sigma_i)\})| \leq 2\sqrt{2\varepsilon} H(\gamma_A(E/\varepsilon)) + 2g(\sqrt{2\varepsilon}) \quad (32)$$

for any ensembles $\{p_i, \rho_i\}$ and $\{q_i, \sigma_i\}$ such that $E(\{p_i, \rho_i\}), E(\{q_i, \sigma_i\}) \leq E$ and $D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \leq \varepsilon \leq \frac{1}{2}$, where γ_A is the Gibbs state of system A .

The metric D_* can be replaced by any of the metrics D_0 and D_K .

Note that the continuity bound (32) does not depend on the channel Φ .

Proof. Condition (20) shows that $H(\gamma_A(E)) = o(\sqrt{E})$, $E \rightarrow +\infty$ (by Lemma 2 in the Appendix). So, continuity bound (32) implies uniform continuity of the function $\{p_i, \rho_i\} \rightarrow \chi(\{p_i, \Phi(\rho_i)\})$ on the set of all ensembles with bounded average energy.

Take arbitrary $\epsilon > 0$. Let $\{\tilde{p}_i, \tilde{\rho}_i\}$ and $\{\tilde{q}_i, \tilde{\sigma}_i\}$ be ensembles belonging respectively to the sets $\mathcal{E}(\{p_i, \rho_i\})$ and $\mathcal{E}(\{q_i, \sigma_i\})$ such that

$$D_*(\{p_i, \rho_i\}, \{q_i, \sigma_i\}) \geq D_0(\{\tilde{p}_i, \tilde{\rho}_i\}, \{\tilde{q}_i, \tilde{\sigma}_i\}) - \epsilon \quad (33)$$

(see the definition (28) of D_*). Consider the qc -states

$$\hat{\rho} = \sum_i \tilde{p}_i \tilde{\rho}_i \otimes |i\rangle\langle i| \quad \text{and} \quad \hat{\sigma} = \sum_i \tilde{q}_i \tilde{\sigma}_i \otimes |i\rangle\langle i|$$

in $\mathfrak{S}(\mathcal{H}_{AC})$, where $\{|i\rangle\}$ is a basis in \mathcal{H}_C . We have

$$\chi(\{p_i, \Phi(\rho_i)\}) = \chi(\{\tilde{p}_i, \Phi(\tilde{\rho}_i)\}) = I(B:C)_{\Phi \otimes \text{Id}_C(\hat{\rho})}$$

and

$$\chi(\{q_i, \Phi(\sigma_i)\}) = \chi(\{\tilde{q}_i, \Phi(\tilde{\sigma}_i)\}) = I(B:C)_{\Phi \otimes \text{Id}_C(\hat{\sigma})}.$$

Since $E(\{p_i, \rho_i\}) = E(\{\tilde{p}_i, \tilde{\rho}_i\}) = \text{Tr} H_A \hat{\rho}_A$ and $E(\{q_i, \sigma_i\}) = E(\{\tilde{q}_i, \tilde{\sigma}_i\}) = \text{Tr} H_A \hat{\sigma}_A$, continuity bound (32) follows from continuity bound (26).

The last assertion of the proposition follows from (29) and (31). \square

3.2.2 Continuous ensembles

In analysis of infinite-dimensional quantum systems and channels the notion of *generalized (continuous) ensemble* defined as a Borel probability measure on the set of quantum states naturally appears [3, 5]. We denote by $\mathcal{P}(\mathcal{H})$ the set of all Borel probability measures on $\mathfrak{S}(\mathcal{H})$ equipped with the topology of weak convergence [2, 12].⁸ The set $\mathcal{P}(\mathcal{H})$ is a complete separable metric space containing the dense subset $\mathcal{P}_0(\mathcal{H})$ of discrete measures (corresponding to discrete ensembles) [2, 12]. The average state of a generalized ensemble $\mu \in \mathcal{P}(\mathcal{H})$ is the barycenter of the measure μ defined by the Bochner integral

$$\bar{\rho}(\mu) = \int_{\mathfrak{S}(\mathcal{H})} \rho \mu(d\rho).$$

For an ensemble $\mu \in \mathcal{P}(\mathcal{H}_A)$ its image $\Phi(\mu)$ under a quantum channel $\Phi : A \rightarrow B$ is defined as the ensemble in $\mathcal{P}(\mathcal{H}_B)$ corresponding to the measure $\mu \circ \Phi^{-1}$ on $\mathfrak{S}(\mathcal{H}_B)$, i.e. $\Phi(\mu)[\mathfrak{S}_B] = \mu[\Phi^{-1}(\mathfrak{S}_B)]$ for any Borel subset $\mathfrak{S}_B \subseteq \mathfrak{S}(\mathcal{H}_B)$, where $\Phi^{-1}(\mathfrak{S}_B)$ is the pre-image of \mathfrak{S}_B under the map Φ . If $\mu = \{\pi_i, \rho_i\}$ then this definition implies $\Phi(\mu) = \{\pi_i, \Phi(\rho_i)\}$.

The output Holevo quantity of a generalized ensemble $\mu \in \mathcal{P}(\mathcal{H})$ is defined as (cf. [5])

$$\chi(\Phi(\mu)) = \int H(\Phi(\rho) \| \Phi(\bar{\rho}(\mu))) \mu(d\rho) = H(\Phi(\bar{\rho}(\mu))) - \int H(\Phi(\rho)) \mu(d\rho),$$

where the second formula is valid under the condition $H(\Phi(\bar{\rho}(\mu))) < +\infty$.

The Kantorovich distance (30) between discrete ensembles is extended to generalized ensembles μ and ν as follows

$$D_K(\mu, \nu) = \frac{1}{2} \inf_{\Lambda \in \Pi(\mu, \nu)} \int_{\mathfrak{S}(\mathcal{H}) \times \mathfrak{S}(\mathcal{H})} \|\rho - \sigma\|_1 \Lambda(d\rho, d\sigma), \quad (34)$$

⁸The weak convergence of a sequence $\{\mu_n\}$ to a measure μ_0 means that $\lim_{n \rightarrow \infty} \int f(\rho) \mu_n(d\rho) = \int f(\rho) \mu_0(d\rho)$ for any continuous bounded function f on $\mathfrak{S}(\mathcal{H})$.

where $\Pi(\mu, \nu)$ is the set of all probability measures on $\mathfrak{S}(\mathcal{H}) \times \mathfrak{S}(\mathcal{H})$ with the marginals μ and ν . Since $\frac{1}{2}\|\rho - \sigma\|_1 \leq 1$ for all ρ and σ , the Kantorovich distance (34) generates the weak convergence topology on $\mathcal{P}(\mathcal{H})$ [2, Ch.8].

For arbitrary generalized ensembles μ and ν there exist sequences $\{\mu_n\}$ and $\{\nu_n\}$ of discrete ensembles converging respectively to μ and ν such that

$$\lim_{n \rightarrow \infty} \chi(\Phi(\mu_n)) = \chi(\Phi(\mu)), \quad \lim_{n \rightarrow \infty} \chi(\Phi(\nu_n)) = \chi(\Phi(\nu))$$

and $\bar{\rho}(\mu_n) = \bar{\rho}(\mu)$, $\bar{\rho}(\nu_n) = \bar{\rho}(\nu)$ for all n . Such sequences can be obtained by using the construction from the proof of Lemma 1 in [5] and taking into account the lower semicontinuity of the function $\mu \mapsto \chi(\Phi(\mu))$ [5, Pr.1]. Since $D_K(\mu_n, \nu_n)$ tends to $D_K(\mu, \nu)$, by using these sequences one can derive (by means of approximation) from Proposition 2 its continuous version.

Proposition 3. *Let $\Phi : A \rightarrow B$ be an arbitrary quantum channel. If the Hamiltonian H_A of input system A satisfies condition (20) then the function $\mu \rightarrow \chi(\Phi(\mu))$ is uniformly continuous on the subset of $\mathcal{P}(\mathcal{H})$ consisting of ensembles μ with bounded average energy $E(\mu) \doteq \text{Tr} H_A \bar{\rho}(\mu)$ and*

$$|\chi(\Phi(\mu)) - \chi(\Phi(\nu))| \leq 2\sqrt{2\varepsilon}H(\gamma_A(E/\varepsilon)) + 2g(\sqrt{2\varepsilon}) \quad (35)$$

for any ensembles μ and ν such that $E(\mu), E(\nu) \leq E$ and $D_K(\mu, \nu) \leq \varepsilon \leq \frac{1}{2}$, where γ_A is the Gibbs state of the system A .

The independence of continuity bound (35) on Φ has several interesting corollaries. One of them is considered in the next section.

3.3 On uniform finite-dimensional approximation of the Holevo capacity of a channel with energy constraint.

In this section we show that speaking about the Holevo capacity of energy constrained infinite-dimensional channels from a given system to any other systems we may consider (permitting arbitrarily small error) that all these channels have *the same finite-dimensional input space* – the subspace corresponding to the minimal eigenvalues of the input Hamiltonian.

The Holevo capacity of the channel $\Phi : A \rightarrow B$ with the (input) energy constraint can be defined as follows:

$$\bar{C}(\Phi, H_A, E) = \sup_{E(\mu) \leq E} \chi(\Phi(\mu)), \quad (36)$$

where the supremum is over all ensembles in $\mathcal{P}(\mathcal{H}_A)$ with the average energy $E(\mu) \doteq \text{Tr} H_A \bar{\rho}(\mu)$ not exceeding E [5]. This quantity determines the ultimate rate of transmission of classical information through the channel Φ by using nonentangled block encoding, for large class of channels it coincides with the classical capacity of Φ under the energy constraint [3, 4].

Assume the Hamiltonian H_A satisfies condition (20). So, it can be represented as follows

$$H_A = \sum_{k=0}^{+\infty} E_k |k\rangle\langle k|,$$

where $E_{k+1} \geq E_k$ and $\{|k\rangle\}$ is the orthonormal basic of eigenvectors of H_A . Let $\mathcal{H}_{H_A}^n$ be the linear span of the vectors $|1\rangle, \dots, |n\rangle$, i.e. $\mathcal{H}_{H_A}^n$ is the eigensubspace of H_A corresponding to its n minimal eigenvalues (taking the multiplicity into account). Let

$$\bar{C}_n(\Phi, H_A, E) = \sup_{E(\mu) \leq E, \text{supp} \mu \subseteq \mathcal{H}_{H_A}^n} \chi(\Phi(\mu)), \quad (37)$$

where the supremum is over all ensembles⁹ in $\mathcal{P}(\mathcal{H}_A)$ supported by the subspace $\mathcal{H}_{H_A}^n$ and such that $E(\mu) \leq E$. The value $\bar{C}_n(\Phi, H_A, E)$ can be treated as the Holevo capacity $\bar{C}(\Phi_n, H_A, E)$ of the restriction $\Phi_n \doteq \Phi|_{\mathfrak{S}(\mathcal{H}_{H_A}^n)}$ of the channel Φ to the set $\mathfrak{S}(\mathcal{H}_{H_A}^n)$.

By using the lower semicontinuity of the function $\mu \mapsto \chi(\Phi(\mu))$ one can show that $\bar{C}_n(\Phi, H_A, E)$ tends to $\bar{C}(\Phi, H_A, E)$ as $n \rightarrow +\infty$ for any given channel Φ . The results of the previous section make it possible to prove that this convergence is *uniform* on the set of all channels from the system A to any systems, i.e. the rate of convergence does not depend on a channel.

Theorem 2. *If the Hamiltonian H_A satisfies condition (20) and $E \geq E_0$ then for any $\varepsilon > 0$ there is natural n_ε such that*

$$0 \leq \bar{C}(\Phi, H_A, E) - \bar{C}_{n_\varepsilon}(\Phi, H_A, E) \leq \varepsilon$$

for arbitrary channel Φ from the system A to any system B .

From the information point of view the above theorem shows that for any given Hamiltonian H_A satisfying condition (20), $E \geq E_0$ and $\varepsilon > 0$ there is n_ε -dimensional subspace $\mathcal{H}_{H_A}^{n_\varepsilon}$ of the input space \mathcal{H}_A such that the Holevo capacity $\bar{C}(\Phi, H_A, E)$ of *any* channel Φ is ε -achievable by nonentangled block encoding used only states supported by the tensor powers of $\mathcal{H}_{H_A}^{n_\varepsilon}$.

⁹The suprema in (36) and in (37) can be taken only over discrete ensembles [5].

Proof. Let $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ be the subset of $\mathcal{P}(\mathcal{H}_A)$ consisting of ensembles μ such that $\text{Tr} H_A \bar{\rho}(\mu) \leq E$ and $\mathcal{P}_{H_A,E}^n(\mathcal{H}_A)$ the subset of $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ consisting of ensembles supported by the subspace $\mathcal{H}_{H_A}^n$.

Since the subset $\{\rho \in \mathfrak{S}(\mathcal{H}_A) \mid \text{Tr} H_A \rho \leq E\}$ is compact [4, the Lemma], the set $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ is compact (in the weak convergence topology) by Proposition 2 in [5]. By using the construction from the proof of Lemma 1 in [5] it is easy to show density in $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ of its subset consisting of discrete ensembles. So, to prove density of the set $\bigcup_n \mathcal{P}_{H_A,E}^n(\mathcal{H}_A)$ in $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ it suffices to show that for any ensemble $\{p_i, \rho_i\}$ in $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ there is a sequence $\{\{p_i^n, \rho_i^n\}\}_n$ converging to $\{p_i, \rho_i\}$ such that $\{p_i^n, \rho_i^n\} \in \mathcal{P}_{H_A,E}^n(\mathcal{H}_A)$ for all n . Such sequence can be constructed as follows. Let P_n be the projector on the subspace $\mathcal{H}_{H_A}^n$, $p_i^n = p_i \text{Tr} P_n \rho_i / \text{Tr} P_n \bar{\rho}$ and $\rho_i^n = P_n \rho_i P_n / \text{Tr} P_n \rho_i$ for all i , where $\bar{\rho}$ is the average state of the ensemble $\{p_i, \rho_i\}$. By using the definition of $\mathcal{H}_{H_A}^n$ it is easy to show that the sequence $\{\{p_i^n, \rho_i^n\}\}_n$ has the required properties (for all n such that $E_n > E$).

Since the set $\mathcal{P}(\mathcal{H}_A)$ can be treated as a metric space with the Kantorovich distance D_K defined in (34), the set $\mathcal{P}_{H_A,E}(\mathcal{H}_A)$ and the sequence $\{\mathcal{P}_{H_A,E}^n(\mathcal{H}_A)\}_n$ of its subsets satisfy the condition of the below Lemma 1. So, this lemma shows that for any $\delta > 0$ there is a natural n_δ such that

$$\inf_{\nu \in \mathcal{P}_{H_A,E}^{n_\delta}(\mathcal{H}_A)} D_K(\mu, \nu) < \delta$$

for any ensemble $\mu \in \mathcal{P}_{H_A,E}(\mathcal{H}_A)$, i.e. in the δ -vicinity of any $\mu \in \mathcal{P}_{H_A,E}(\mathcal{H}_A)$ there is an ensemble $\nu \in \mathcal{P}_{H_A,E}^{n_\delta}(\mathcal{H}_A)$.

Condition (20) guarantees that for any $\varepsilon > 0$ we can choose such δ that

$$2\sqrt{2\delta}H(\gamma_A(E/\delta)) + 2g(\sqrt{2\delta}) < \varepsilon,$$

where γ_A is the Gibbs state of the system A corresponding to the energy E . So, the assertion of the theorem follows directly from Proposition 3 (and definitions (36) and (37)). \square

Lemma 1. *Let \mathcal{K} be a compact subset of a metric space with the metric D and $\{\mathcal{K}_n\}$ a sequence of subsets of \mathcal{K} such that $\mathcal{K}_n \subseteq \mathcal{K}_{n+1}$ and $\bigcup_n \mathcal{K}_n$ is dense in \mathcal{K} . Then*

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathcal{K}} \inf_{y \in \mathcal{K}_n} D(x, y) = 0.$$

Proof. Assume for any n there is $x_n \in \mathcal{K}$ such that $D(x_n, y) \geq \delta > 0$ for all $y \in \mathcal{K}_n$. By the compactness of \mathcal{K} there is a subsequence $\{x_{n_k}\}$ converging to some $x_* \in \mathcal{K}$. Then the assumed property of $\{x_n\}$ implies that the

$\delta/3$ -vicinity of x_* has empty intersection with all the sets \mathcal{K}_n contradicting to the density of $\bigcup_n \mathcal{K}_n$ in \mathcal{K} . \square

In [16] it is proved that the "unconstrained" Holevo capacity is uniformly continuous on the set of all channels with given finite-dimensional input space with respect to the diamond norm

$$\|\Psi\|_\diamond \doteq \sup_{\rho \in \mathfrak{T}(\mathcal{H}_{AR}), \|\rho\|_1=1} \|\Psi \otimes \text{Id}_R(\rho)\|_1,$$

coinciding with the norm of complete boundedness of the dual map Ψ^* to the map Ψ [3, 18].

Similar arguments show that the same property holds for the Holevo capacity of constrained channels. So, Theorem 2 implies the following

Corollary 5. *If the Hamiltonian H_A satisfies condition (20) then the function*

$$\Phi \mapsto \bar{C}(\Phi, H_A, E)$$

is uniformly continuous on the set of all channels from the system A to any system B in the diamond norm topology.

Appendix: auxiliary lemmas

Lemma 2. *Condition (20) implies that*

$$\sup_{\text{Tr} H_B \rho < E} H(\rho) = o(\sqrt{E}), \quad E \rightarrow +\infty.$$

Proof. Condition (20) shows that $\text{Tr} e^{-\lambda H_B} < +\infty$ for all $\lambda > 0$. So, the operator H_B has the discrete spectrum $\{E_k\}_{k \geq 0}$, where we assume that $E_{k+1} \geq E_k$ for all k . Condition (20) means that

$$\lim_{\lambda \rightarrow +0} \lambda g(\lambda) = 0, \quad \text{where} \quad g(\lambda) = \log \sum_{k=0}^{+\infty} e^{-\lambda E_k}. \quad (38)$$

Let $f(E) \doteq \sup_{\text{Tr} H_B \rho < E} H(\rho)$. It is shown in the proof of Proposition 1 in [13] that $f'(E) = \lambda(E)$ for all $E \in [E_0, +\infty)$, where $\lambda(E)$ is a differentiable strictly decreasing function determined by the equality

$$\sum_{k=0}^{+\infty} E_k e^{-\lambda E_k} = E \sum_{k=0}^{+\infty} e^{-\lambda E_k} \quad (39)$$

such that

$$\lim_{E \rightarrow E_0+0} \lambda(E) = +\infty \quad \text{and} \quad \lim_{E \rightarrow +\infty} \lambda(E) = 0. \quad (40)$$

By L'Hopital's rule to prove that $f(E) = o(\sqrt{E})$ it suffices to show that

$$\lim_{E \rightarrow +\infty} \sqrt{E} \lambda(E) = 0. \quad (41)$$

Denote by $E(\lambda)$ the inverse function to $\lambda(E)$. Equality (39) implies that

$$E(\lambda) = -g'(\lambda), \quad (42)$$

where $g(\lambda)$ is the function defined in (38). It follows from (40) and (42) that (41) can be rewritten as

$$\lim_{\lambda \rightarrow +0} \lambda^2 g'(\lambda) = 0. \quad (43)$$

So, to prove the lemma it suffices to show that (38) implies (43). Assume that (43) is not valid. Then there exists a vanishing sequence $\{\lambda_n\}$ of positive numbers such that $\lambda_n^2 |g'(\lambda_n)| \geq \delta > 0$ for all n . Since (42) and the strict concavity of $f(E)$ imply that

$$g''(\lambda) = -E'(\lambda) = -1/\lambda'(E) = -1/f''(E) > 0,$$

the positive function $g(\lambda)$ is convex. It follows that for any λ_n and $\lambda \in (0, \lambda_n)$ we have

$$g(\lambda) \geq g(\lambda_n) + |g'(\lambda_n)|(\lambda_n - \lambda) \geq g(\lambda_n) + \delta(\lambda_n - \lambda)/\lambda_n^2$$

and hence

$$\lambda g(\lambda) \geq \lambda g(\lambda_n) + \delta \lambda (\lambda_n - \lambda) / \lambda_n^2 \geq \delta \lambda (\lambda_n - \lambda) / \lambda_n^2.$$

By taking $\lambda = \lambda_n/2$ we obtain $(\lambda_n/2)g(\lambda_n/2) \geq \delta/4$ for all n contradicting to (38). \square

Lemma 3. *Let $E_k = \log^q k$, $k = 1, 2, \dots$, then $\lim_{\lambda \rightarrow +0} [\sum_{k \geq 1} e^{-\lambda E_k}]^\lambda = 1$ if and only if $q > 2$.*

Proof. Note that $\sum_{k \geq 1} e^{-\lambda E_k} < +\infty$ for all $\lambda > 0$ if and only if $q > 1$.

For any $q > 1$ we have

$$\int_1^{+\infty} e^{-\lambda \log^q x} dx \leq \sum_{k=1}^{+\infty} e^{-\lambda E_k} \leq \int_1^{+\infty} e^{-\lambda \log^q x} dx + 1. \quad (44)$$

By introducing the variable $u = \lambda^{1/q} \log x$ we obtain

$$I(\lambda) \doteq \int_1^{+\infty} e^{-\lambda \log^q x} dx = \lambda^{-1/q} \int_0^{+\infty} e^{-u^q + u\lambda^{-1/q}} du.$$

If $q > 2$ then

$$\int_0^1 e^{-u^q + u\lambda^{-1/q}} du \leq \int_0^1 e^{u\lambda^{-1/q}} du = \lambda^{1/q} [e^{\lambda^{-1/q}} - 1]$$

and

$$\begin{aligned} & \int_1^{+\infty} e^{-u^q + u\lambda^{-1/q}} du \leq \int_1^{+\infty} e^{-u^2 + u\lambda^{-1/q}} du \\ &= \int_1^{+\infty} e^{-(u-0.5\lambda^{-1/q})^2 + 0.25\lambda^{-2/q}} du \leq e^{0.25\lambda^{-2/q}} \int_{-\infty}^{+\infty} e^{-t^2} dt = \sqrt{\pi} e^{0.25\lambda^{-2/q}}. \end{aligned}$$

Since $2/q < 1$, these estimates show that $\lim_{\lambda \rightarrow +0} \lambda \log I(\lambda) = 0$. Hence the right inequality in (44) implies $\lim_{\lambda \rightarrow +0} [\sum_k e^{-\lambda E_k}]^\lambda = 1$ in this case.

If $q = 2$ then

$$\begin{aligned} I(\lambda) &= \lambda^{-1/2} \int_0^{+\infty} e^{-u^2 + u\lambda^{-1/2}} du = \lambda^{-1/2} \int_0^{+\infty} e^{-(u-0.5\lambda^{-1/2})^2 + 0.25\lambda^{-1}} du \\ &\geq \lambda^{-1/2} e^{0.25\lambda^{-1}} \int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2} \lambda^{-1/2} e^{0.25\lambda^{-1}}. \end{aligned}$$

So, in this case $\lim_{\lambda \rightarrow +0} \lambda \log I(\lambda) \neq 0$ and the left inequality in (44) implies $\lim_{\lambda \rightarrow +0} [\sum_k e^{-\lambda E_k}]^\lambda \neq 1$. \square

Acknowledgments. I am grateful to A.S.Holevo for useful remarks. I am also grateful to the participants of the workshop "Recent advances in continuous variable quantum information theory", Barcelona, April, 2016 (especially to A.Winter) for stimulating discussion.

This work is supported by the Russian Science Foundation under grant 14-21-00162.

References

- [1] R.Alicki, M.Fannes, "Continuity of quantum conditional information", Journal of Physics A: Mathematical and General, V.37, N.5, L55-L57 (2004); arXiv: quant-ph/0312081.
- [2] V.I.Bogachev, "Measure theory", Springer-Verlag Berlin Heidelberg, 2007.
- [3] A.S.Holevo, "Quantum systems, channels, information. A mathematical introduction", Berlin, DeGruyter, 2012.
- [4] A.S.Holevo, "Classical capacities of quantum channels with constrained inputs", Probability Theory and Applications. V.48, N.2, 359-374 (2003); arXiv:quant-ph/0211170.
- [5] A.S.Holevo, M.E.Shirokov, "Continuous ensembles and the χ -capacity of infinite dimensional channels", Theory of Probability and Applications, V.50, N.1, 86-98 (2005); arXiv:quant-ph/0408176.
- [6] A.A.Kuznetsova, "Quantum conditional entropy for infinite-dimensional systems", Theory of Probability and its Applications, V.55, N.4, 709-717 (2011).
- [7] G.Lindblad, "Expectation and Entropy Inequalities for Finite Quantum Systems", Comm. Math. Phys., V.39, N.2, 111-119 (1974).
- [8] G.Lindblad, "Entropy, information and quantum measurements", Comm. Math. Phys., V.33, 305-322 (1973).
- [9] N.Li, S.Luo, "Classical and quantum correlative capacities of quantum systems", Phys. Rev. A 84, 042124 (2011).
- [10] M.A.Nielsen, I.L.Chuang, "Quantum Computation and Quantum Information", Cambridge University Press, 2000.
- [11] O.Oreshkov, J.Calsamiglia, "Distinguishability measures between ensembles of quantum states", Phys. Rev. A 79, 032336 (2009); arXiv:0812.3238.
- [12] K.R.Parthasarathy, "Probability measures on metric spaces", Academic Press, New York and London, 1967.

- [13] M.E.Shirokov "Entropic characteristics of subsets of states", *Izvestiya: Mathematics*, 70:6, 1265-1292 (2006); quant-ph/0510073.
- [14] M.E.Shirokov, "Measures of correlations in infinite-dimensional quantum systems", *Sbornik: Mathematics*, 207:5, 724-768 (2016); arXiv:1506.06377.
- [15] M.E.Shirokov, "Tight continuity bounds for the quantum conditional mutual information, for the Holevo quantity and for capacities of a channel", arXiv:1512.09047 (v.2!).
- [16] M.E.Shirokov, "Continuity bounds for information characteristics of quantum channels depending on input dimension", arXiv:1604.00568.
- [17] A.Wehrh, "General properties of entropy", *Rev. Mod. Phys.* V.50, 221-250 (1978).
- [18] M.M. Wilde, "From Classical to Quantum Shannon Theory", arXiv:1106.1445 (v.6).
- [19] M.M. Wilde, H.Qi, "Energy-constrained private and quantum capacities of quantum channels", arXiv:1609.01997.
- [20] A.Winter, "Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints", *Comm. Math. Phys.*, V.347 N.1, 291-313 (2016); arXiv:1507.07775 (v.6).